

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 7, 2005, 07:04:17 ; Search time 8.94985 Seconds

(without alignments)  
1193.323 Million cell updates/sec

Title: US-09-939-537-37

Perfect score: 591

Sequence: 1 TRFSSABPAPYQGGQGNOLY.....LSTATKDTYDALHMQALPPR 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	583	98.6	163	2 A31768	T-cell receptor ze
2	476.5	80.6	164	2 A40104	T-cell receptor CD
3	471.5	79.8	166	2 JC4664	T-cell receptor ze
4	471.5	79.8	166	2 A46424	T-cell surface gly
5	389	65.8	188	2 A45089	T-cell receptor CD
6	378.5	64.0	177	2 S54817	T-cell receptor co
7	378.5	64.0	206	2 A35900	T-cell receptor CD
8	86	14.6	1154	2 S69206	regulator protein
9	80	13.5	692	2 T21095	hypothetical prote
10	76.5	12.9	527	2 S38329	probable ATP-depen
11	74	12.5	136	2 T14305	probable late embr
12	74	12.5	3190	2 T13828	CREB-binding prote
13	73.5	12.4	123	1 S26183	general stress pro
14	73.5	12.4	360	2 AF0317	rare lipoprotein A
15	73.5	12.4	374	2 H70415	conserved hypochet
16	73.5	12.4	1258	2 T30522	nuclear protein SA
17	71.5	12.1	632	2 T02627	hypothetical prote
18	70.5	11.9	594	2 G96525	protein T1N15.25 (
19	70	11.8	193	2 AG2567	hypothetical prote
20	70	11.8	448	2 A46311	70K UI small nucle
21	70	11.8	568	1 I61106	involucrin - rat
22	70	11.8	2441	2 D71633	erythrocyte membra
23	69.5	11.8	185	2 S71512	hypothetical prote
24	69.5	11.8	205	2 T20962	hypothetical prote
25	69.5	11.8	1269	2 F84730	probable myosin he
26	69	11.7	393	2 T05532	hypothetical prote
27	69	11.7	488	2 G81213	conserved hypochet
28	69	11.7	490	2 C81790	conserved hypochet
29	68.5	11.6	615	2 G87678	ATP-dependent DNA

30	68	11.5	123	2 S07364	seminal vesicle se
31	68	11.5	251	2 T17784	PBCV-1 33kd peptid
32	68	11.5	319	2 I50519	sox 19 protein - z
33	68	11.5	384	1 A43710	involucrin - chick
34	67.5	11.4	212	2 S40027	protein p71 - mous
35	67.5	11.4	249	2 P95957	probable transcrip
36	67.5	11.4	761	2 F82078	topoisomerase IV,
37	67.5	11.4	784	2 G95112	exoribonuclease,
38	67.5	11.4	836	2 C97525	clpA protein (A722
39	67.5	11.4	836	2 AD2744	ATP-dependent Clp
40	67.5	11.4	963	2 T04002	hypothetical prote
41	67	11.3	140	2 T40368	hypothetical prote
42	67	11.3	247	2 AB3638	probable lipoprote
43	67	11.3	438	2 B84633	peptide chain rele
44	67	11.3	677	1 S09078	chromogranin B pre
45	67	11.3	898	2 T25167	hypothetical prote

#### ALIGNMENTS

RESULT 1  
A31768  
T-cell receptor zeta chain precursor - human  
C/Species: Homo sapiens (man)  
C/Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
C/Accession: A31768  
R/Weisman, A.M.; Hou, D.; Orloff, D.G.; Modi, W.S.; Senauer, H.; O'Brien, S.J.; Klausner, J.; Weissman, A.M.; U.S.A. 85, 9709-9713, 1988  
A/Title: Molecular cloning and chromosomal localization of the human T-cell receptor zeta  
A/Reference number: A31768; MUID:89071765; PMID:2974162  
A/Accession: A31768  
A/Molecule type: mRNA  
A/Residues: 1-163 <MBI>  
A/Cross-references: UNIPROT:P20963; GB:J04132; NID:G623041; PIDN:AAA60394.1; PID:G623042  
C/Keywords: phosphoprotein, T-cell receptor, transmembrane protein  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-163/Product: T-cell receptor zeta chain #status predicted <MAT>

Query Match 98.6%; Score 583; DB 2; Length 163;  
Best Local Similarity 99.1%; Pred. No. 1.9e-49;  
Matches 109; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	2	RFSRSABPAPYQGGQGNOLYNELNGRREYDVLDKRGGRDPENGGKPRRKNPOGLYNEL	61
Db	54	KFSRSABPAPYQGGQGNOLYNELNGRREYDVLDKRGGRDPENGGKPRRKNPOGLYNEL	113

Qy 62 QKDMAEAVSEIGMKGRGKGDGLYQGLSTATDTYDALHMQALPPR 111  
Db 114 QKDMAEAVSEIGMKGRGKGDGLYQGLSTATDTYDALHMQALPPR 163

RESULT 2  
A40104  
T-cell receptor CD3 zeta chain - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 09-Jul-2004  
C/Accession: A40104; I55293  
R/Weisman, A.M.; Baniyash, M.; Hou, D.; Samelson, L.E.; Burgess, W.H.; Klausner, R.D.  
A/Title: Molecular cloning of the zeta chain of the T cell antigen receptor.  
A/Reference number: A40104; MUID:88145643; PMID:3278377  
A/Accession: A40104  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-164 <MBI>  
A/Cross-references: UNIPROT:P24161; GB:M19729; NID:G201131; PIDN:AAA40171.1; PID:G201132  
R/Baniyash, M.; Hsu, V.W.; Seidlin, M.F.; Klausner, R.D.  
J. Biol. Chem. 264, 13252-13257, 1989  
A/Title: The isolation and characterization of the murine T cell antigen receptor zeta c  
A/Reference number: I55293; MUID:89337299; PMID:2787796  
A/Accession: I55293  
A/Status: preliminary; translated from GB/EMBL/DBJ



QY 2 RFRSAEPYAOQGNOLYNELNGRREYVDLDRGRDPENWGK-PRRKNPOEGLYNE 60  
Db 54 KFRSAETAANLQDPNOLYNELNGRREYVDLEKGRARDPENWGKQRRRNPOEGLYNA 113  
QY 61 LQDKMAEAYSEIGMGERRRGKHGDLQY 90  
Db 114 LQDKMAEAYSEIGTGERRRRGKHGDLQY 143

## RESULT 7

A35900  
T-cell receptor CD3 eta chain precursor - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 23-Oct-1990 #sequence\_revision 23-Oct-1990 #text\_change 09-Jul-2004  
C/Accession: A35900; A60374; G66522  
R/Jin, Y.J.; Clayton, L.K.; Howard, F.D.; Koyasu, S.; Sieh, M.; Steindrich, R.; Tarr, G.  
Proc. Natl. Acad. Sci. U.S.A. 87, 3319-3323, 1990  
A/Title: Molecular cloning of the CD3eta subunit identifies a CD3zeta-related product in  
A/Reference number: A35900; MUID:90239005; PMID:2139725  
A/Accession: A35900  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-206 <IN>  
A/Cross-references: UNIPROT:P29020; GB:M33158; NID:G192488; PIDN:AAA37398.1; PID:G309155  
R/Ohno, H.; Saito, T.  
Int. Immunol. 2, 1117-1119, 1990  
A/Title: CD3zeta and eta chains are produced by alternative splicing from a common gene.  
A/Reference number: A60374; MUID:91190781; PMID:2150596  
A/Accession: A60374  
A/Status: not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 144-206 <OH>  
R/Jensen, J.P.; Cencelarelli, C.; Hou, D.; Relihan, B.L.; Dean, M.; Weissman, A.M.  
J. Immunol. 150, 122-130, 1993  
A/Title: T cell antigen receptor-eta subunit. Low levels of expression and limited cross  
A/Reference number: A46522; MUID:93107707; PMID:8417118  
A/Contents: annotation  
A/Comment: The functional significance of this alternatively spliced product of the CD3  
lnc acids, differs widely among various mammalian species in sequence, length, and even  
C/Keywords: alternative splicing; T-cell receptor; transmembrane protein

Query Match 64.0%; Score 378.5; DB 2; Length 206;  
Best Local Similarity 82.2%; Pred. No. 1,7e-29;  
Matches 74; Conservative 5; Mismatches 10; Indels 1; Gaps 1;  
QY 2 RFRSAEPYAOQGNOLYNELNGRREYVDLDRGRDPENWGK-PRRKNPOEGLYNE 60  
Db 54 KFRSAETAANLQDPNOLYNELNGRREYVDLEKGRARDPENWGKQRRRNPOEGLYNA 113  
QY 61 LQDKMAEAYSEIGMGERRRGKHGDLQY 90  
Db 114 LQDKMAEAYSEIGTGERRRRGKHGDLQY 143

## RESULT 8

S69206  
regulator protein white collar 1 - Neurospora crassa  
C/Species: Neurospora crassa  
C/Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 15-Mar-2004  
C/Accession: S69206  
R/Ballario, P.; Vitorioso, P.; Magrelli, A.; Talora, C.; Cabibbo, A.; Macino, G.  
EMBO J. 15, 1650-1657, 1996  
A/Title: White collar-1, a central regulator of blue light responses in Neurospora, is a  
A/Reference number: S69206; MUID:96203083; PMID:8612589  
A/Accession: S69206  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1154 <BAL>  
A/Cross-references: EMBL:X94300; NID:G1279576; PID:G1480115  
C/Genetics:  
A/Intons: 967/3  
C/Keywords: zinc finger  
F/332-991/Domain: GATA-type zinc finger homology <GF>

Query Match 14.6%; Score 86; DB 2; Length 1154;  
Best Local Similarity 24.8%; Pred. No. 2.8;  
Matches 33; Conservative 19; Mismatches 49; Indels 32; Gaps 6;

QY 9 PPAYOQGNOLYNELNGRREYVDLDRGRDPENWGKRR-XXXKN 52  
Db 71 PPTTNGNSTHNSDVMSGSDSLDEIQNLDEHRRSVPPYGGTRRLSMFDYAN 130  
QY 53 PDEGLYNELQDKMAEAYSE-IGMGERRRGK-----HGLVQGLS---TATK 97  
Db 131 PNDG-FSDVQLDMSNGYDGTGGMGSHSPYAGNTAMSDHSGYSHMSPNVMGNN 189  
QY 98 DTYDALHMQALPP 110  
Db 190 MTPPLNMTMHSPP 202

## RESULT 9

T21095  
hypothetical protein F18H3.3a - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T21095  
R/Coles, L.  
submitted to the EMBL Data Library, July 1995  
A/Reference number: Z19373  
A/Accession: T21095  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-692 <ML>  
A/Cross-references: UNIPROT:Q19579; EMBL:Z50110; PIDN:CAA90444.1; GSPDB:GN00028; CESP:FLI  
A/Experimental source: clone F18H3  
C/Genetics:  
A/Gene: CESP:F18H3.3a  
A/Map position: X  
A/Intons: 111/1; 215/2; 469/3; 552/3  
C/Superfamily: polyadenylate-binding protein; ribonucleoprotein repeat homology

Query Match 13.5%; Score 80; DB 2; Length 692;  
Best Local Similarity 21.7%; Pred. No. 6;  
Matches 30; Conservative 22; Mismatches 40; Indels 46; Gaps 6;  
QY 9 PPAYOQGNOLYNELNGRREYVDLDRGRDP----- 42  
Db 486 PGOQVPOQNFQYPPNQRYT---PQGRPRMTNDGRPOYGAPRPGVGGPG 542  
QY 43 -EMGKPRRN-----PDEGLYNELQDKMAEAYSEIGMGERRRGKHGDL-YQGLS 93  
Db 543 VQNGAPRINQGGAPRPGAPQKPRYQGRPQGRQGRPAQGPQGGQSGSIVIHGE 602  
QY 94 TATKDYDALHM--QALP 109  
Db 603 TLTS-----HMLAQAP 614

## RESULT 10

S38329  
probable ATP-dependent RNA helicase Dbp45A - fruit fly (Drosophila melanogaster)  
N/Alternate names: DEAD-box protein  
C/Species: Drosophila melanogaster  
C/Date: 31-Dec-1993 #sequence\_revision 02-Aug-1994 #text\_change 09-Jul-2004  
C/Accession: S38329; S38749  
R/Lavoie, C.A.; Harvey, M.; Laeko, P.F.  
Biochim. Biophys. Acta 1216, 140-144, 1993  
A/Title: Dbp45A encodes a Drosophila DEAD box protein with similarity to a putative yeast  
A/Reference number: S38329; MUID:94032476; PMID:7692973  
A/Accession: S38329  
A/Molecule type: DNA  
A/Residues: 1-527 <LAVI>  
A/Cross-references: UNIPROT:Q07886; GB:L13612  
A/Note: the authors translated the codon TGC for residue 211 as Ser, GAG for residue 212  
4 as Gln, GGT for residue 225 as Arg, GTG for residue 228 as Leu, TCG for residue 229 as



[illegible]

Db 140 LVPHHDLVAGSVGYSALIKENNEYDTV 168

Search completed: March 7, 2005, 07:21:03

Job time : 11.1499 secs

```

RESULT 14
AF0317
rare lipoprotein A precursor [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AF0317
R:Parkhill, J.; Wen, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:1156360
A:Accession: AF0317
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <KUR>
A:Cross-references: UNIPROT:Q8ZDG6; GB:AL590842; PIDD:QAC92845.1; PID:q15960589; GSPDB:G
C:Genetics:
A:Gene: rlpA
A:Superfamily: rlpA_lipoprotein

```

```
Query Match      12.4% Score 73.5; DB 2, Length 360;  
Beet Local Similarity 25.9%, Pred. No. 12,  
Matches 29, Conservative 17, Mismatches 33, Indels 33, Gaps 6,  
  
QY    8 EPPAYOQGQNLTN---ELNLGRRRBEVDVLDKRGGRPEWGGKKPR--KNPOEGLYNEL 61  
       :||| |::: ||: |: ::::: ||: |:||  
Db     21 QPAPQQOOVOOTYGAVVEIG-GABPRYEFPNPVNQYKVNGOSYRIIKPKPM----- 73  
               * * * * *  
QY    62 QKDRAEAYSSEIGMK---GERRRGKHGLQLGSSTATDRTDYDLAHMALPP 110  
       :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db     74 -----FSDIGLASSYGEARGN-----TTATGELTFDPNALPAHP 108
```

RESULT 15  
H70415  
conserved hypothetical protein [aQ\\_1336](#) - *Aquifex aeolicus*  
C/Species: *Aquifex aeolicus*  
C/Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C/Accession: [H70415](#)  
R/Dickert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.B.; O  
V.  
Nature 392, 353-358, 1998  
A/Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
A/Reference number: A70300; MUID:98196666; PMID:9537320  
A/Accession: [H70415](#)  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-374 <AOF>  
A/Cross-references: UNIPROT:O67355; GB:AE000735; NID:g2983749; PIDN:AAO07322.1; PID:g2983749  
A/Experimental source: strain VF5  
C/Genetics:  
A/Gene: [aQ\\_1336](#)

Query Match	12.4%	Score	73.5	DB	2	Length	374
Best Local	29.2%	Pred.	No. 13				
Matches	26	Conservative	21	Mismatches	37	Indels	5
						Gaps	4

```
QY 19 LVNELNLGR-REBYDVLDRGRDPEMGKPPRKPOSGLVNELQK--DKMAEAYSELGM 75
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 81 LVYERFLNTLKKERREKILTKGYRPESHGAQVAPENPOE-LKNPIETVKNKSEKFKARGI 139
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 76 KGRRRGKGHDGLYQGLSTATQDT--VDAL 103
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

**This Page Blank (uspto)**